DOCUMENT RESUME

ED 090 297 TM 003 582

AUTHOR Cooper, James M.
TITLE Pupil Growth Measures and Teacher Evaluation: Pressing for Practical Applications.
PUB DATE Apr 74
NOTE 19p.; Paper presented at the American Educational Research Association Annual Meeting (Chicago, Illinois, April 15-19, 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; Effective Teaching; Feedback; *Performance Based Teacher Education; Performance Tests; *Preservice Education; *Student Development; Student Teaching; Teacher Behavior; Teacher Education; *Teacher Evaluation; Teaching Skills

ABSTRACT

The use of pupil growth measures to evaluate the competency of preservice teacher trainees is advocated by many proponents of competency-based teacher education. This paper (1) raises pro and con arguments for the use of pupil growth measures to evaluate preservice teachers, (2) suggests practical applications for the use of various types of pupil growth measures to evaluate teacher trainees, and (3) suggests ways in which these pupil growth measures can be developed and measured. (Author)

ED 090297

TM 003 582

27.08

A PAPER PRESENTED AT THE ANNUAL MEETING OF THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, CHICAGO, ILLINOIS, APRIL 15-19, 1974

PUPIL GROWTH MEASURES AND TEACHER EVALUATION: PRESSING FOR PRACTICAL APPLICATIONS

JAMES M. COOPER
UNIVERSITY OF HOUSTON

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THERE IS LITTLE DOUBT THAT THE MOST DOMINANT TREND IN TEACHER EDUCATION TODAY IS THE COMPETENCY BASED TEACHER EDUCATION MOVEMENT. WHILE THIS MOVEMENT LACKS PRECISE DEFINITION MOST PROPONENTS OF CBTE WOULD AGREE THAT ONE OF ITS UNDERLYING ASSUMPTIONS IS THAT THE DEMONSTRATION OF SPECIFIED TEACHER BEHAVIORS AND THE TEACHER'S ABILITY TO BRING ABOUT INTENDED LEARNER OUTCOMES ARE MORE VALID MEASURES OF A TEACHER'S COMPETENCE THAN PRESAGE VARIABLES THAT ONLY MEASURE A TEACHER'S KNOWLEDGE. WHILE MANY PROCESS VARIABLES (TEACHER BEHAVIORS) HAVE BEEN ADVOCATED BY TEACHER EDUCATORS AS BEING DESIRABLE, THE RELATIONSHIPS BETWEEN THESE TEACHER BEHAVIORS AND INTENDED PUPIL OUTCOMES ARE, AT BEST, SKETCHY (ROSENSHINE AND FURST, 1971). NUMEROUS RESEARCHERS AND TEACHER EDUCATORS BELIEVE THAT THE APPLICATION OF PUPIL PRODUCT CRITERIA TO DETERMINE TEACHER EFFECTIVENESS IS LESS INFERENTIAL AND MORE PREDICTIVE THAN THE APPLICATION OF EITHER TEACHER-PUPIL PROCESS CRITERIA OR TEACHER PRESAGE CRITERIA (MCNEIL AND POPHAM, 1973; SCHALOCK, 1971; COOPER AND WEBER, 1973). WHILE THIS THEORETICAL POSITION IS HELD BY MANY EDUCATORS, THE PRACTICAL APPLICATON OF PUPIL GROWTH MEASURES TO EVALUATE PRESERVICE TEACHER TRAINEES HAS RARELY BEEN MADE. THIS PAPER WILL SUGGEST SOME PRACTICAL APPLICATIONS FOR THE USE OF PUPIL GROWTH MEASURES TO EVALUATE TEACHER TRAINEES.

Rationale For Using Pupil Outcome Data to Measure Preservice Teacher Effectiveness

Researchers have agreed for years that the ultimate criterion of a teacher's competence is his ability to bring about desired pupil outcomes (AERA, 1952; Biddle and Ellena, 1964; McNeil and Popham, 1973). Some advantages of this position have been outlined by Schalock:

> 1) It represents or provides an absolute criterion of teaching effectiveness....
>
> 2) It accommodates individual differences in teaching preferences or styles in that it allows for wide variation in the means of teaching a given outcome....
>
> 3) It allows for the fact that at this point in time we are not at all clear about the specific teaching behaviors that bring about specified outcomes in pupils, or the specific behaviors that bring about selected noninstructional outcomes; but it does require that effective behaviors and/or instructional programs be found and utilized.
>
> 4) It forces the entire educational system, including teacher education, to be clear about the goals or objectives of education, and to become clear about the means for the realization of those objectives (Schalock, 1971).

While theoretical agreement regarding the value of using pupil outcomes to measure teacher effectiveness is easily obtained, the practicality of the issue has been seriously questioned by many researchers (Mitzel, 1960; Lawler, 1964; Flounders, 1965; Smith, 1967; and Soar, 1973). Among their concerns are: (1) the adequacy of measures for assessing a wide range of pupil outcomes in

DIFFERENT SUBJECT-MATTER AREAS AND AT DIFFERENT EDUCATIONAL LEVELS; (2) DETERMINING WHAT ROLE THE TEACHER PLAYED, AS OPPOSED TO OTHER INSTRUCTIONAL VARIABLES, IN PROMOTING THE DESIRED PUPIL OUTCOMES; (3) MEASUREMENT AND STATISTICAL PROBLEMS, INCLUDING REGRESSION AND CEILING EFFECTS; (4) THE ASSESSMENT OF RELATIVELY COMPLEX PROBLEM SOLVING AND RESPONSIBLE CITIZENSHIP BEHAVIOR WHICH MAY NOT BE MEASURABLE WITHIN A LIMITED TIME PERIOD; AND (5) DIFFERENCES IN LEARNING APTITUDES AMONG PUPILS WHICH MAY MAKE THE TEACHER EFFECTIVE WITH SOME CHILDREN AND NOT WITH OTHERS.

ALL OF THE ABOVE CONCERNS ARE LEGITIMATE AND VALID. THE AUTHOR KNOWS OF NO RESPONSIBLE EDUCATIONAL RESEARCHER OR TEACHER EDUCATOR WHO ARGUES FOR RELYING SOLELY ON MEASURES OF PUPIL OUTCOMES TO EVALUATE INDIVIDUAL TEACHER EFFECTIVENESS FOR THE VERY REASONS CITED ABOVE. DESPITE THESE PROBLEMS, THE RECENT DEVELOPMENT OF CRITERION-REFERENCED MEASUREMENT, TEACHING PERFORMANCE TESTS, AND CONTRACT PLANS LEAD THE AUTHOR TO BELIEVE THAT SOME MEASURES OF PUPIL OUTCOMES CAN BE USED TO HELP EVALUATE TEACHERS IN BOTH A FORMATIVE AND SUMMATIVE SENSE.

THE POTENTIAL FOR USING PUPIL OUTCOME DATA TO HELP EVALUATE PRESERVICE TEACHERS IS MUCH MORE LIMITED, AND CERTAINLY NOT AS CONTROVERSIAL, THAN THEIR POSSIBLE USE WITH INSERVICE TEACHERS. TO BEGIN WITH, THE TEACHER TRAINEE DOES NOT HAVE CONSTANT EXPOSURE TO, NOR FULL RESPONSIBILITY FOR, GROUPS OF CHILDREN. WHEREAS THE REGULAR CLASSROOM TEACHER TEACHES ALL DAY FOR 9 MONTHS, THE TEACHER

TRAINEE IS USUALLY LIMITED TO ONE SEMESTER OR LESS OF STUDENT TEACHING, ONLY A PORTION OF WHICH DOES HE HAVE FULL RESPONSIBILITY FOR THE CLASS, POSSIBLY PRECEDED BY SOME TUTORING AND MICROTEACHING. THIS RESTRICTED EXPOSURE TO CHILDREN DURING THE TEACHER EDUCATION PROGRAM NECESSARILY LIMITS THE KINDS OF PUPIL OUTCOMES THAT THE TRAINEE CAN REASONABLY BE EXPECTED TO ACHIEVE. AT THE PRESERVICE STAGE OF TEACHING THE TYPES OF PUPIL OUTCOMES THAT THE TRAINEE IS EXPECTED TO DEMONSTRATE SHOULD PROBABLY BE VIEWED IN TERMS OF MINIMUMS RATHER THAN MAXIMUMS. THAT IS, FOR INITIAL CERTIFICATION THE TRAINEE SHOULD BE EXPECTED TO BRING ABOUT RELATIVELY SIMPLE TYPES OF PUPIL OUTCOMES, WITHIN LIMITED TIME PERIODS. THE REST OF THE PAPER WILL EXPLORE THE KINDS OF PUPIL OUTCOMES WHICH MIGHT BE EXPECTED, HOW THEY MIGHT BE ASSESSED, THE CONTEXTS IN WHICH INSTRUCTION MIGHT OCCUR, AND WHO SHOULD DETERMINE THE DESIRED PUPIL OUTCOMES.

Teaching Performance Tests

RECENTLY EDUCATIONAL RESEARCHERS HAVE BEEN EXPERIMENTING WITH THE USE OF TEACHING PERFORMANCE TESTS AS A MEASUREMENT APPROACH DESIGNED TO ASSESS TEACHING COMPETENCY. (POPHAM, 1973; MCNEIL AND POPHAM, 1973) BRIEFLY, TEACHING PERFORMANCE TESTS FUNCTION AS FOLLOWS: THE TEACHER IS GIVEN A SET OF EXPLICIT INSTRUCTIONAL OBJECTIVES (PLUS A SAMPLE TEST ITEM) AND IS ASKED TO PREPARE A SHORT LESSON DESIGNED TO HELP THE STUDENTS ACHIEVE THE OBJECTIVES. IF THE TEACHER IS NOT FAMILIAR WITH THE TOPIC, RELEVANT BACKGROUND MATERIAL IS GIVEN TO HIM. THE TEACHER PLANS THE LESSON AND THEN INSTRUCTS A GROUP OF LEARNERS (EITHER ADULTS OR CHILDREN) FOR A GIVEN PERIOD OF TIME, TYPICALLY FIFTEEN TO THIRTY MINUTES. THE

NUMBER OF LEARNERS MAY VARY FROM A FEW TO AN ENTIRE CLASS.

AT THE END OF THE INSTRUCTIONAL PERIOD THE LEARNERS ARE GIVEN A POSTTEST BASED ON THE OBJECTIVES. A PRETEST MAY OR MAY NOT BE GIVEN TO THE LEARNERS DEPENDING ON WHETHER THE SUBJECT MATTER IS LIKELY TO BE KNOWN BY SOME OF THEM. THE POSTTEST, WHICH HAS NOT BEEN SEEN PREVIOUSLY BY THE TEACHER, REFLECTS THE OBJECTIVES. THE LEARNERS ARE ALSO ASKED TO RATE HOW INTERESTING THE LESSON WAS. THE TEACHER IS APPRISED THAT THIS RATING WILL OCCUR AND IS NOT ONLY ASKED TO ACCOMPLISH THE OBJECTIVE BUT TO PLAN A LESSON THAT LEARNERS WILL FIND INTERESTING. THUS, AN ESTIMATE OF THE TEACHER'S ABILITY TO PROMOTE PRESPECIFIED OBJECTIVES IS OBTAINED FROM THESE TWO INDICATORS.

TEACHING PERFORMANCE TESTS THUS FOCUS ON TWO MEASURES OF LEARNER OUTCOMES: (1) PERFORMANCE ON A TEST DESIGNED TO MEASURE THE PRE-SPECIFIED OBJECTIVES OF THE LESSON (USUALLY COGNITIVE IN NATURE), AND 2) INTEREST LEVEL OF THE LEARNERS AS MEASURED BY THEIR RATINGS (AN AFFECTIVE TYPE OF MEASUREMENT).

CAN TEACHING PERFORMANCE TESTS BE USED IN A PRESERVICE PROGRAM AS AN INDICATOR OF THE TEACHER'S POTENTIAL EFFECTIVENESS? THAT IS, CAN ONE ARGUE THAT THOSE PROSPECTIVE TEACHERS WHO EXPERIENCE THE GREATEST SUCCESS ON TEACHING PERFORMANCE TESTS ARE MORE APT TO PROMOTE DESIRED LEARNER OUTCOMES IN REGULAR TEACHING SITUATIONS? IT IS TEMPTING TO RESPOND AFFIRMATIVELY ON THE BASIS OF PRIMA FACIE EVIDENCE. HOWEVER, A NUMBER OF RESEARCHERS (GLASS, 1972; POPHAM, 1973; AND BAKER, 1973) ARGUE AGAINST JUMPING TO THIS CONCLUSION SO EARLY IN THE HISTORY OF TEACHING PERFORMANCE TESTS. POPHAM (1973) CAUTIONS THAT "DEVELOPMENTAL WORK WITH TEACHING PERFORMANCE TESTS IS STILL AT SUCH AN EARLY STAGE THAT IT MAY BE IMPRUDENT TO EMPLOY THEM FOR THE EVALUATION OF <u>INDIVIDUAL</u>

TEACHERS. THE ONLY EXCEPTION MIGHT BE FOR ISOLATING INSTRUCTORS WHO ARE EXTREMELY WEAK OR STRONG IN THEIR ABILITY TO ACCOMPLISH PRESPECIFIED GOALS." BAKER (1973) WARNS THAT IF PERFORMANCE TESTS ARE TO BE USED FOR DECISION PURPOSES, SUCH AS THE SELECTION OR EVALUATION OF TEACHERS, THE ISSUES OF CONSISTENCY AND VALIDITY MUST BE ADDRESSED TO A GREATER EXTENT THAN HAS OCCURRED THUS FAR IN THE HISTORY OF PERFORMANCE TESTS. SHE SUGGESTS THAT PERFORMANCE TESTS FOR TEACHER EVALUATION COULD BE USED TO IDENTIFY THE INDIVIDUAL WHO HAS BEEN GIVEN TIME AND ASSISTANCE AND IS STILL UNABLE TO DEMONSTRATE INFLUENCE OVER THE OUTCOMES OF INSTRUCTION.

MCNEIL AND POPHAM (1973) STATE THAT THE RELIABILITY FOR DETERMINING TEACHING COMPETENCY BY USING PERFORMANCE TESTS CAN BE INCREASED BY USING A NUMBER OF LESSONS, DIFFERENT KINDS OF OBJECTIVES AND DIFFERENT SUBJECT MATTER. THEY ALSO CITE SEVERAL STUDIES WHICH SUPPORT THE CONCLUSION THAT WHEN THERE IS REASONABLE CONTROL OVER EXTRANEOUS FACTORS SUCH AS TEACHER FAMILIARITY WITH CONTENT AND PUPIL POPULATIONS, SOME TEACHERS ARE CONSISTENTLY MORE SUCCESSFUL THAN OTHER IN GETTING INTENDED RESULTS WITH PUPILS.

IF ONE ACCEPTS THAT THE ULTIMATE CRITERION OF A TEACHER'S EFFECTIVENESS IS BEST MEASURED BY EXAMINING DESIRED PUPIL OUTCOMES, THEN THE USE OF TEACHING PERFORMANCE TESTS IS A MAJOR BREAKTHROUGH IN THE EVALUATION OF PRESERVICE TEACHERS. NO ONE CAN ARGUE WITH THE FACT THAT TEACHER TRAINEES ARE EVALUATED BASED ON WHAT IS BELIEVED TO BE EITHER THEIR POTENTIAL OR DEMONSTRATED EFFECTIVENESS AS A TEACHER. THE DATA THAT ARE TYPICALLY USED IN THIS EVALUATION PROCESS ARE: 1) MEASURES OF THEIR KNOWLEDGE AND UNDERSTANDING OF CONCEPTS AND PRINCIPLES RELATED TO INSTRUCTIONAL METHODOLOGIES, LEARNING, AND SUBJECT MATTER CONTENT; AND 2) TEACHING PERFORMANCES AND

BEHAVIORS THAT ARE DEMONSTRATED IN MICRO OR ACTUAL CLASSROOM CONTEXTS. RARELY ARE PUPIL OUTCOMES USED AS A MEASURE OF POTENTIAL TEACHER EFFECTIVENESS; ONLY RECENTLY HAVE A FEW TEACHER EDUCATION PROGRAMS USED PUPIL OUTCOME DATA AS INDICATORS OF COMPETENCY.

UNTIL NOW TEACHING PERFORMANCE TESTS HAVE BEEN DEVELOPED WITHOUT ANY OVERALL SYSTEM GOVERNING THEIR CONSTRUCTION. TOPICS AND OBJECTIVES HAVE BEEN SELECTED STRICTLY ON THE BASIS OF THE DEVELOPERS' WISHES AND INTERESTS. THIS APPROACH IS CLEARLY INADEQUATE IF THE PUPIL DATA COLLECTED FROM TEACHING PERFORMANCE TESTS ARE TO BE USED BOTH AS FEEDBACK FOR TEACHER GROWTH AND IN HELPING TO JUDGE POTENTIAL TEACHER EFFECTIVENESS. THE AUTHOR SUGGESTS THAT A NUMBER OF VARIABLES, IN ADDITION TO THOSE RELATED TO TEST CONSTRUCTION, MUST BE CONSIDERED IN THE FUTURE DEVELOPMENT OF TEACHING PERFORMANCE TESTS. THESE INCLUDE:

1) TYPES OF OBJECTIVES (COGNITIVE, AFFECTIVE)
2) SUBJECT MATTER CONTENT
3) TIME REQUIRED TO TEACH THE LESSONS
4) NUMBERS AND TYPES OF STUDENTS

TYPES OF OBJECTIVES

COGNITIVE. THE LESSONS TO BE TAUGHT AND THE OBJECTIVES DERIVED FROM THOSE LESSONS SHOULD BE DESIGNED TO REFLECT THE VARIOUS TAXONOMIC LEVELS OF THE COGNITIVE DOMAIN. TEACHER TRAINEES SHOULD FIRST TEACH LESSONS WHERE THE PUPIL OBJECTIVES FOCUS ON DEMONSTRATION OF KNOWLEDGE, GRADUALLY MOVING TO LESSONS WHOSE OBJECTIVES REFLECTED THE UPPER LEVELS OF THE TAXONOMY. IN THIS WAY TRAINEES CAN DISCOVER THAT DIFFERENT TYPES OF TEACHING STRATEGIES ARE REQUIRED FOR PUPIL ACHIEVEMENT OF DIFFERENT TYPES OF OBJECTIVES. THE TYPES OF OBJECTIVES SELECTED ARE ALSO AFFECTED BY THE OTHER VARIABLES TO BE DISCUSSED,

NOTABLY SUBJECT MATTER CONTENT AND TIME.

AFFECTIVE. BECAUSE OF THE DIFFICULTY IN MEASURING THE ACHIEVEMENT OF AFFECTIVE OBJECTIVES AND OFTEN THE TIME REQUIRED TO EFFECT CHANGE, THERE WILL PROBABLY BE RELATIVELY FEW AFFECTIVE OBJECTIVES THAT TEACHER TRAINEES WOULD BE EXPECTED TO ACHIEVE WITH PUPILS. THIS IS NOT TO DENY THE IMPORTANCE OF SUCH OBJECTIVES, BUT RATHER TO RECOGNIZE THE LIMITS OF TEACHER EDUCATION PROGRAMS WITH REGARD TO THIS AREA. HOWEVER, THERE ARE SOME TYPES OF AFFECTIVE OBJECTIVES LIKE THOSE RELATED TO VALUE CLARIFICATION WHICH PRESERVICE TEACHERS COULD BE EXPECTED TO ACHIEVE WITH PUPILS. LESSONS OF THIS SORT SHOULD BE DEVELOPED, RECOGNIZING THAT THE TIME REQUIRED TO ACHIEVE OBJECTIVES OF THIS TYPE WILL BE GREATER THAN THAT REQUIRED FOR MANY COGNITIVE OBJECTIVES.

SUBJECT MATTER CONTENT

MOST TEACHING PERFORMANCE TESTS HAVE BEEN DEVELOPED AROUND TOPICS WHICH WERE SUFFICIENTLY ESOTERIC THAT THE NEED FOR A PRETEST WAS NOT THOUGHT NECESSARY. SINCE THE CHOSEN TOPICS WERE RELATIVELY OBSCURE THERE WAS LITTLE CHANCE THAT THE TEACHER WOULD POSSESS MUCH ADVANCE KNOWLEDGE OF THE TOPIC, AN IMPORTANT CONSIDERATION FOR RESEARCH PURPOSES. HOWEVER, WHAT IS A SOUND RESEARCH PRINCIPLE IS, IN THIS INSTANCE, A TEACHER EVALUATION WEAKNESS. WHEN A TEACHER TRAINEE'S COMPETENCE IS BEING ASSESSED, HIS KNOWLEDGE OF THE SUBJECT MATTER FIELDS BEING TAUGHT IN THE SCHOOLS SHOULD NOT BE CONTROLLED FOR, BUT RATHER SHOULD BE INCORPORATED INTO THE ASSESSMENT SCHEMA. THEREFORE THE AUTHOR BELIEVES THAT TEACHING PERFORMANCE TESTS SHOULD BE DEVELOPED BASED ON CURRICULA CURRENTLY BEING TAUGHT IN THE PUBLIC SCHOOLS.

USING A PRETEST-POSTTEST EVALUATION PROCESS, TEACHING PERFORMANCE TESTS COULD BE DEVELOPED ON SUCH CURRICULA AS BIOLOGICAL SCIENCES CURRICULUM STUDY (BSCS), SCHOOL MATHEMATICS STUDY GROUP (SMSG),

Science--A Process Approach, and Directed Instructional Systems for Teaching Arithmetic and Reading (DISTAR).

Having to teach and achieve prespecified objectives with pupils in curriculum areas that are a regular part of their curriculum adds a validity dimension to the teaching performance tests that has heretofore generally been lacking. If teaching performance tests are developed based on content found in public school curricula, actual school-age children must be used. In other words, the lessons to be taught must be matched with children of proper age and ability levels, and who through pretest performances have shown that they have not already achieved the objectives.

Time Required To Teach the Lessons

Most teaching performance tests have been of short duration, typically 15-30 minutes in length, and focusing on one or two pupil objectives. While some kinds of objectives can be taught in this short period of time, many others cannot. Short mini-lessons should be used initially with preservice teachers until the teacher has demonstrated some level of competence in helping the learners achieve the prespecified objectives designed for these mini-lessons. Longer units of learning could then be introduced. These longer units of learning would contain more objectives and/or more complex objectives that were sequentially ordered. This longer teaching performance test would require the teacher trainee to plan and teach units of learning that more closely relate to actual instruction in the schools than do the mini-lessons. These units might range from three to ten lessons of 15-30 minutes in length, and might be taught in a microteaching or actual school

CONTEXT. A SIMILAR PROCESS WAS EMPLOYED DURING THE MID-1960'S IN THE MICROTEACHING CLINIC OPERATED AT STANFORD UNIVERSITY. ALTHOUGH THE STANFORD MICROTEACHING CLINIC FOCUSED ON TEACHING SKILLS RATHER THAN HELPING PUPILS ACHIEVE PRESPECIFIED OBJECTIVES, THE FIRST THREE WEEKS WERE DEVOTED TO MINI-LESSONS WHILE THE LAST THREE WEEKS REQUIRED THE INTERN TEACHERS TO PLAN AND TEACH A UNIT OF INSTRUCTION LASTING TWELVE DAYS, 25-30 MINUTES PER DAY (COOPER AND STROUD, 1967).

THESE LONGER UNITS OF LEARNING MIGHT ALSO BE TAUGHT IN SCHOOLS DURING THE STUDENT TEACHING PERIOD. IF TEACHING PERFORMANCE TESTS WERE DEVELOPED THAT WERE BASED ON CURRICULA BEING USED IN THE SCHOOLS THEN THE STUDENT TEACHER COULD TEACH THE UNITS OF LEARNING AS A REGULAR PART OF SOME CHILDREN'S CURRICULUM. AS TEACHER EDUCATION PROGRAMS BECOME MORE FIELD-ORIENTED AND CLOSE RELATIONSHIPS WITH THE SCHOOLS ARE ESTABLISHED, THIS TYPE OF STUDENT TEACHER EVALUATION BECOMES POSSIBLE.

NUMBERS AND TYPES OF STUDENTS

MUCH OF THE EXPERIENCE WITH TEACHING PERFORMANCE TESTS TO DATE HAS BEEN WITH THE USE OF PEER TEACHING RATHER THAN WITH ACTUAL SCHOOL-AGE CHILDREN. THE REASONS FOR THIS ARE OBVIOUS TO ANYONE WHO HAS EVER TRIED TO IMPLEMENT A MICROTEACHING LABORATORY; IT IS DIFFICULT TO USE CHILDREN UNLESS THEY CAN BE PAID FOR THEIR SERVICES AND THE LABORATORY CAN BE OPERATED DURING NON-SCHOOL HOURS. THE FACT THAT PEERS WERE USED AS PUPILS (ALTHOUGH NOT IN A ROLE-PLAYING SENSE) ALSO HELPS TO EXPLAIN WHY SO MANY LESSONS DEALING WITH RATHER ESOTERIC SUBJECTS HAVE BEEN DEVELOPED INSTEAD OF LESSONS BASED ON ACTUAL ELEMENTARY OR SECONDARY SCHOOL CURRICULA. THE AUTHOR CONTENDS, HOWEVER, THAT UNLESS CHILDREN REPRESENTATIVE OF THE AGE GROUP THE TEACHER TRAINEE INTENDS TO TEACH ARE USED IN THE TEACHING PERFORMANCE TESTS,

THE VALIDITY OF THE EXERCISE IS TOO QUESTIONABLE EVER TO USE FOR SUMMATIVE EVALUATION PURPOSES. PERFORMANCE WITH PEER TEACHING IS EASILY DISCOUNTED BY TEACHERS WHO ARGUE THAT IT IS SO DIFFERENT FROM TEACHING CHILDREN THAT ONE CANNOT GENERALIZE. NOR IS THE AUTHOR FAMILIAR WITH EVIDENCE THAT WOULD SUPPORT THE CONTENTION THAT THOSE WHO ARE SUCCESSFUL IN TEACHING THEIR PEERS ARE SUCCESSFUL IN TEACHING CHILDREN. DESPITE THE LOGISTICAL DIFFICULTIES ENCOUNTERED IN USING ACTUAL CHILDREN, IF TEACHING PERFORMANCE TESTS ARE TO HAVE FACE VALIDITY WITH THOSE WHO ARE MOST DIRECTLY AFFECTED, THE TEACHER TRAINEES, IT IS IMPERATIVE THAT CHILDREN RATHER THAN PEERS BE USED.

THE NUMBER OF STUDENTS TO BE TAUGHT IN ANY TEACHING PERFORMANCE IS ARBITRARY. TWO FACTORS SHOULD BE CONSIDERED, HOWEVER: (1) AS THE NUMBER OF STUDENTS INCREASES, THE TEACHING TASK BECOMES MORE COMPLEX; AND (2) WHEN THE NUMBER OF STUDENTS TAUGHT APPROXIMATES THE NUMBER THAT WOULD BE TAUGHT IN AN ACTUAL CLASSROOM SITUATION, ONE CAN HAVE MORE CONFIDENCE IN THE RESULTS.

INITIAL MINI-LESSONS SHOULD PROBABLY BE TAUGHT TO SMALL (1-6 CHILDREN) GROUPS IN ORDER TO HELP REDUCE THE COMPLEXITY OF THE TASK. AS THE TRAINEE DEMONSTRATES SUCCESS, THE NUMBER OF PUPILS TAUGHT SHOULD BE INCREASED WHEN IT IS FEASIBLE TO DO SO. SOME TYPES OF OBJECTIVES ON THE OTHER HAND, WILL LEND THEMSELVES TO A TUTORING SITUATION WHERE THE TRAINEE WILL WORK WITH ONE CHILD OVER A PERIOD OF TIME. MANY CONDITIONS WILL GOVERN THE NUMBER OF STUDENTS USED SO IT IS DIFFICULT TO BE VERY PRESCRIPTIVE IN THIS PAPER.

CONSIDERATION OF THE TWO FACTORS MENTIONED ABOVE SHOULD GUIDE TEACHER EDUCATORS IN MAKING THESE DECISIONS.

IT WOULD BE NICE TO SAY THAT EACH TRAINEE SHOULD BE SYSTEMATICALLY EXPOSED TO STUDENTS WITH DIFFERENT TYPES OF LEARNING APTITUDES IN

ORDER TO DEMONSTRATE SUCCESS IN ACHIEVING PRESPECIFIED OBJECTIVES WITH CHILDREN POSSESSING DIFFERENT APTITUDES. SINCE WE ARE NOT SURE WHAT APTITUDES ARE RELEVANT, OR HOW THEY CAN BE ASSESSED, THE ISSUE BECOMES A RESEARCH RATHER THAN A TRAINING QUESTION AT THIS POINT IN TIME. HOWEVER, STUDENTS CAN BE GROUPED ACCORDING TO A NUMBER OF CHARACTERISTICS IF IT IS DEEMED DESIRABLE TO DO SO.

FIGURE I GRAPHICALLY SUMMARIZES THE PRECEDING DISCUSSION ON TEACHING PERFORMANCE TEST VARIABLES. WITHIN THE LIMITS OF THE CONSTRAINTS PLACED UPON THEM, TEACHER EDUCATORS SHOULD STRIVE TO UTILIZE THE MORE COMPLEX/AUTHENTIC VARIABLES WHEN CONSTRUCTING TEACHING PERFORMANCE TESTS.

TEACHER SPECIFIED PUPIL OUTCOMES

ONE CHARACTERISTIC OF TEACHING PERFORMANCE TESTS IS THAT THE OBJECTIVES FOR THE LESSONS ARE PRESPECIFIED BY THE DEVELOPERS. WHILE THIS FEATURE HAS THE ADVANTAGE OF COMPARING DIFFERENT TEACHERS' ABILITY TO BRING ABOUT CERTAIN OBJECTIVES WITH COMPARABLE SETS OF STUDENTS, THE FACT THAT THE OBJECTIVES ARE PRESPECIFIED DOES NOT ALLOW THE TEACHER THE OPPORTUNITY TO DEVELOP OBJECTIVES BASED UPON THE BACKGROUND, KNOWLEDGE AND INTEREST LEVEL OF THE STUDENTS. TEACHING PERFORMANCE TESTS DO NOT ALLOW THE TEACHER TO DEMONSTRATE HIS JUDGMENT CONCERNING THE APPROPRIATENESS AND ATTAINABILITY OF OBJECTIVES FOR A GIVEN SET OF LEARNERS. FOR THIS REASON, THE AUTHOR BELIEVES THAT IN ADDITION TO PUPIL OUTCOME DATA OBTAINED FROM TEACHING PERFORMANCE TESTS, PUPIL DATA FROM LESSONS IN WHICH THE TEACHER DETERMINES THE OBJECTIVES SHOULD ALSO BE CONSIDERED.

ONE STRATEGY FOR OBTAINING PUPIL OUTCOME DATE FROM LESSONS IN WHICH THE TEACHER DETERMINES BOTH THE OBJECTIVES AND THE LEARNING

TEACHING PERFORMANCE TEST VARIABLES

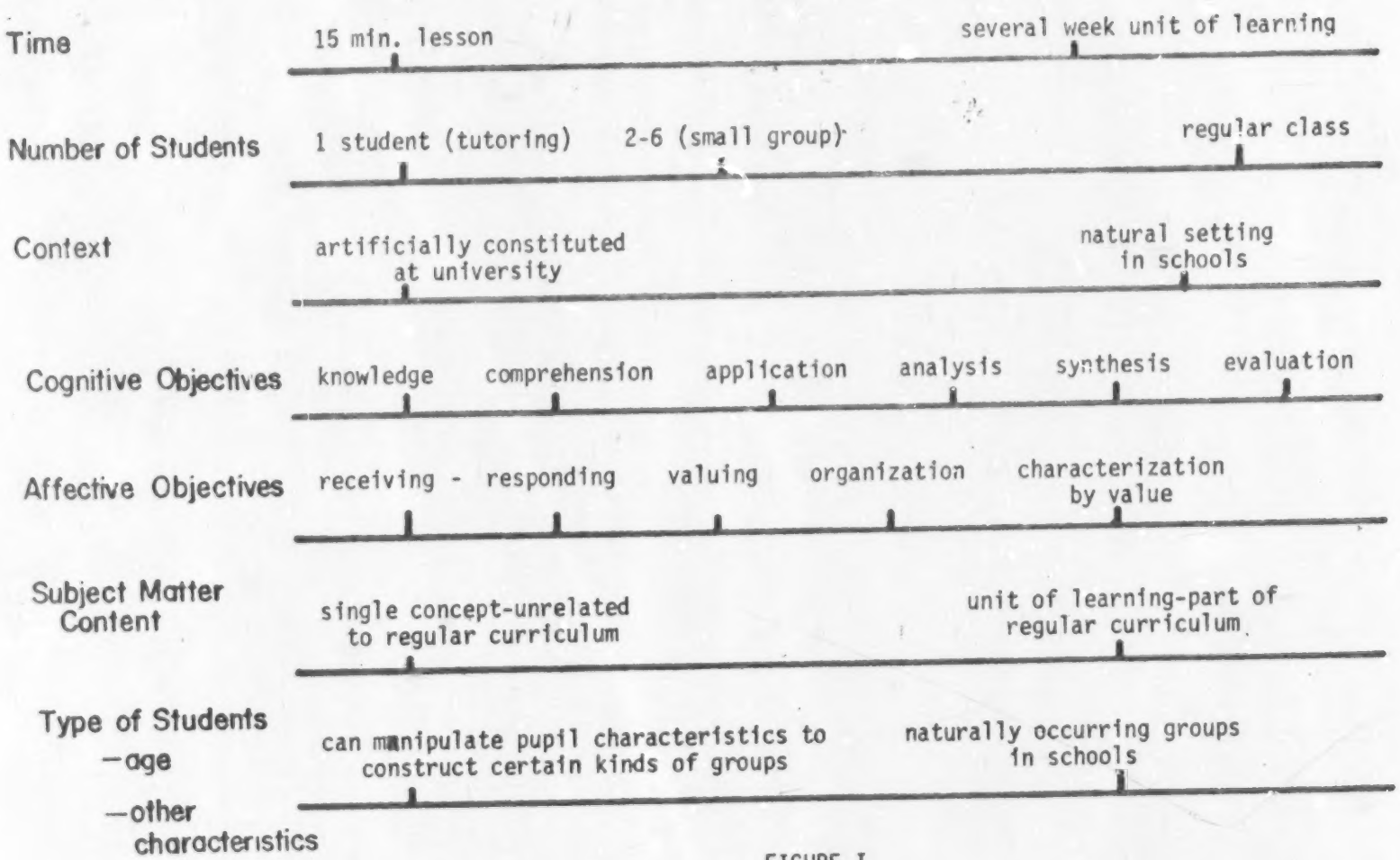

FIGURE I

Variables increase in complexity and/or authenticity from left to right.

ACTIVITIES IS THE CONTRACT PLAN (McNEIL AND POPHAM, 1963). THEY DESCRIBE THE CONTRACT PLAN AS FOLLOWS:

> THE ESSENCE OF THIS TECHNIQUE INVOLVES THE DEVELOPMENT OF A CAREFULLY SELECTED SET OF OBJECTIVES FOR THE PUPIL. SUPERVISORS AND TEACHERS AGREE IN ADVANCE WHAT THEY WILL ACCEPT AS EVIDENCE THAT THE TEACHER HAS BEEN SUCCESSFUL IN CHANGING THE SKILLS, COMPETENCIES OR ATTITUDES OF HIS STUDENTS. AN AGREEMENT IS DRAWN UP BEFORE THE TEACHER INSTRUCTS AND IS DESIGNED TO COUNTER THE PREVAILING PRACTICE OF TRYING TO MAKE AN EX POST FACTO JUDGMENT ABOUT THE DESIRABILITY OF ENDS. SUBSEQUENTLY, EVIDENCE IS COLLECTED TO SEE HOW WELL THE LEARNERS ACHIEVED THE STATED OBJECTIVES AS WELL AS WHETHER UNINTENDED OUTCOMES HAVE EMERGED....CONTRACTS ARE PREPARED FOR VARYING PERIODS OF TIME-A SINGLE DAY'S LESSON, A SEMESTER PLAN, A YEAR OF INSTRUCTION. THIS CONTRACT SYSTEM DEMANDS THAT DATA BY WHICH TO JUDGE MORE CLEARLY WHAT THE INSTRUCTION HAS DONE TO THOSE WHO HAVE BEEN SUBJECTED TO IT BE SUPPLIED, AND, WHEN COUPLED WITH INSTRUCTIONAL ANALYSES, OUGHT TO ENABLE A TEACHER TO REVISE AND BETTER IN SOME RESPECT THE PROCEDURES EMPLOYED IN PREVIOUS WORK (McNEIL AND POPHAM, 1973, PAGE 234).

THE CONTRACT PLAN APPEARS TO BE VIABLE FOR BOTH PRESERVICE AND INSERVICE TEACHERS. AT THE PRESERVICE LEVEL THE CONTRACTS WOULD DEAL WITH OBJECTIVES THAT COULD BE ATTAINABLE WITHIN A TIME PERIOD IN WHICH THE TEACHER TRAINEE HAD ACCESS TO LEARNERS. IT SHOULD BE NOTED THAT IF THE TEACHER SUCCEEDS IN HAVING AN AGREED UPON PERCENTAGE OF HIS STUDENTS ACHIEVE THE OBJECTIVES THAT HE SPECIFIED, THE LEVEL OF COMPETENCY THAT HE HAS DEMONSTRATED IS SITUATION SPECIFIC AND NOT GENERALIZABLE TO OTHER CONTEXTS, CHILDREN, OR OBJECTIVES. HOWEVER, THE USE OF THE PUPIL OUTCOME DATA IN THIS SPECIFIC SITUATION IS STILL IMPORTANT IN MAKING EITHER FORMATIVE OR SUMMATIVE EVALUATIONS OF THE TEACHER. THE AUTHOR WOULD FAR PREFER TO HAVE EVIDENCE OF PUPIL

GROWTH FROM ONE SPECIFIC SITUATION THAN NOT TO HAVE ANY EVIDENCE OF PUPIL GROWTH AT ALL.

AT LEAST ONE TEACHER EDUCATION PROGRAM WITH WHICH THE AUTHOR IS FAMILIAR REQUIRES ITS TRAINEES TO DEMONSTRATE THAT THEY CAN BRING ABOUT DESIRED LEARNING OUTCOMES IN THEIR PUPILS BEFORE THEY CAN BE CERTIFIED. THE OREGON COLLEGE OF EDUCATION, ELEMENTARY TEACHER EDUCATION PROGRAM, REQUIRES THAT ITS TRAINEES ASSUME FULL RESPONSIBILITY FOR THE LEARNING OF PUPILS THROUGHOUT A TWO TO FIVE WEEK PERIOD OF TIME IN ORDER TO BE CERTIFIED. DURING THIS TIME PERIOD THE TRAINEES MUST PLAN AND PREPARE FOR INSTRUCTION, PERFORM INSTRUCTIONAL FUNCTIONS, BRING ABOUT DESIRED LEARNING OUTCOMES IN PUPILS, RELATE INTERPERSONALLY, AND CARRY OUT RELATED PROFESSIONAL RESPONSIBILITIES. THE COMPETENCIES, AND STANDARDS FOR JUDGING THE COMPETENCIES, ARE SPECIFIED IN ADVANCE. THE PERFORMANCE STANDARDS USED IN RELATION TO THE COMPETENCY CLUSTER, "BRINGING ABOUT DESIRED LEARNING OUTCOMES IN PUPILS," ARE AS FOLLOWS:

STANDARD 1. AT LEAST SIXTY PERCENT OF THE PUPILS TAUGHT WILL ACHIEVE THE LEARNING OUTCOMES EXPECTED FROM THE UNIT OF INSTRUCTION PREPARED ESPECIALLY FOR THE 2 TO 5 WEEK TEACHING EXPERIENCE, AND AT LEAST AN ADDITIONAL TWENTY PERCENT WILL SHOW EVIDENCE OF APPRECIABLE PROGRESS IN RELATION TO THOSE OUTCOMES. THIS STANDARD MUST BE DEMONSTRATED FOR EACH SUBJECT MATTER AREA INCLUDED IN THE UNIT AND FOR EACH CLASS OF LEARNING OUTCOME INTENDED FROM THE UNIT, FOR EXAMPLE, KNOWELDGE, SKILL, ATTITUDE.

STANDARD 2. TWO OF THE THREE CHILDREN WORKED WITH INDIVIDUALLY DURING THE COURSE OF THE 2 TO 5 WEEK TEACHING EXPERIENCE SHALL DEMONSTRATE EVIDENCE OF GAIN IN OVERCOMING THE READING PROBLEM(S) IDENTIFIED AT THE ONSET OF THE 2 TO 5 WEEK TEACHING EXPERIENCE.

STANDARD 3. SOME PROGRESS IN THE USE OF SELECTED PROBLEM SOLVING AND SOCIAL INTERACTION SKILLS CAN BE SHOWN FOR THE CLASS AS A WHOLE (SCHALOCK AND FERGUSON, 2-5-74, APPENDIX B).

THE UNIT OF INSTRUCTION FOR THE 2 TO 5 WEEK TEACHING EXPERIENCE, INCLUDING A DESCRIPTION OF THE LEARNING OUTCOMES, THE INDICATORS TO BE USED AS EVIDENCE OF THE REALIZATION OF THOSE OUTCOMES, AND AN OUTLINE OF THE INSTRUCTIONAL MATERIALS AND PROCEDURES TO BE USED, HAVE TO BE APPROVED BY BOTH THE UNIVERISTY SUPERVISOR AND THE COOPERATING TEACHER. BOTH OF THEM ALSO EVALUATE THE TRAINEE'S PERFORMANCE LEVEL WITH RESPECT TO THE STANDARDS. SHOULD A TRAINEE FAIL TO BRING ABOUT THE INTENDED LEARNING OUTCOMES WITH THE PUPILS, HE WOULD BE GIVEN AN OPPORTUNITY TO TRY AGAIN. IF AFTER SEVERAL ATTEMPTS THE TRAINEE IS UNABLE TO SUCCEED, EITHER WITH THE SAME OR OTHER GROUPS OF PUPILS, HE WILL NOT BE RECOMMENDED FOR CERTIFICATION. IT SHOULD BE NOTED THAT EACH TRAINEE HAD TO SUCCESSFULLY MEET THE PERFORMANCE STANDARDS (WHICH DID NOT INCLUDE ACTUALLY BRINGING ABOUT DESIRED PUPIL OUTCOMES) IN TEACHING 2-5 LESSONS AND IN TEACHING 2-5 DAYS BEFORE BEING ALLOWED TO TRY TEACHING 2-5 WEEKS.

SUMMARY

TEACHER TRAINEES IN MOST TEACHER EDUCATION PROGRAMS ARE EVALUATED PRIMARILY ON THE BASIS OF KNOWLEDGE THEY HAVE ACQUIRED AND TO A LESSER EXTENT ON DEMONSTRATED TEACHING SKILLS. THESE PRESAGE AND PROCESS CRITERIA USED TO EVALUATE TRAINEES GENERALLY INCLUDE PAPER-AND-PENCIL TESTS AND TERM PAPERS, EVALUATIONS FROM MICRO-TEACHING SESSIONS, AND STUDENT TEACHING OBSERVATIONS OF THE COOPERATING TEACHER AND THE COLLEGE SUPERVISOR. ONLY RARELY ARE PUPIL OUTCOMES USED EITHER FOR FORMATIVE FEEDBACK OR AS A MEASURE OF POTENTIAL EFFECTIVENESS. THE AUTHOR BELIEVES THAT SOME KINDS OF PUPIL GROWTH DATA CAN AND SHOULD BE USED TO EVALUATE PERSERVICE TEACHERS.

TEACHING PERFORMANCE TESTS WHICH CONTAIN PRESPECIFIED OBJECTIVES ARE A RECENT PHENOMENON AND MUCH WORK STILL REMAINS BEFORE AN ADEQUATE POOL OF SUCH TESTS IS AVAILABLE, AND IN WHICH TEACHER EDUCATORS CAN HAVE CONFIDENCE. IF A POOL OF TEACHING PERFORMANCE TESTS BECOMES LARGE AND DIVERSE ENOUGH TO TEST DIFFERENT TYPES OF LEARNINGS, IT MAY BE THAT THEY WILL BECOME A GOOD PREDICTOR OF POTENTIAL EFFECTIVENESS. IF, DURING THE COURSE OF A TEACHER'S TRAINING, HE WERE REQUIRED TO TEACH MANY DIFFERENT LESSONS, WITH DIFFERENT TYPES OF OBJECTIVES AND SUBJECT MATTER CONTENT, TO PUPILS REPRESENTATIVE OF THOSE HE INTENDS TO TEACH, HE WOULD DEVELOP A PROFILE OF HIS TEACHING EFFECTIVENESS. TEACHER EDUCATORS COULD HAVE MORE CONFIDENCE IN THIS PROFILE AS A PREDICTOR OF FUTURE SUCCESS THAN IN MOST CRITERIA THAT ARE CURRENTLY USED.

IN ADDITION, PRESERVICE TEACHERS SHOULD BE GIVEN OPPORTUNITIES TO SPECIFY THEIR OWN OBJECTIVES THAT THEY WISH TO TEACH TO STUDENTS IN SCHOOLS. PUPIL OUTCOME DATA FROM THESE TEACHER-SPECIFIED OBJECTIVES CAN BE USED IN A FORMATIVE SENSE TO HELP THE TEACHER IMPROVE INSTRUCTIONAL PROCESSES. PUPIL OUTCOME DATA CAN ALSO BE USED IN A SUMMATIVE SENSE TO HELP SUPERVISORS MAKE DECISIONS ABOUT TRAINEES REGARDING RECOMMENDATIONS FOR CERTIFICATION. USING PUPIL OUTCOME DATA AS DESCRIBED IN THIS PAPER, WOULD CERTAINLY INCREASE THE AUTHOR'S CONFIDENCE THAT THE TEACHERS PRODUCED FROM SUCH A PROGRAM WOULD BE ABLE TO BRING ABOUT DESIRED PUPIL OUTCOMES IN THEIR CLASSES. AFTER ALL, THAT'S THE NAME OF THE GAME.

REFERENCES

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, COMMITTEE ON THE CRITERIA OF TEACHER EFFECTIVENESS. "REPORT OF THE COMMITTEE ON THE CRITERIA OF TEACHER EFFECTIVENESS." REVIEW OF EDUCATIONAL RESEARCH, 1952. 22, 238-263.

BAKER, EVA L. "TEACHING PERFORMANCE TESTS AS DEPENDENT MEASURES IN INSTRUCTIONAL RESEARCH." PAPER PRESENTED AT THE ANNUAL MEETING OF AREA, NEW ORLEANS, 1973.

COOPER, JAMES M. AND WEBER, WILFORD A. "WHO HAS THE POTENTIAL TO BECOME AN EFFECTIVE TEACHER?" IN KEITH GOLDHAMMER (ED.), PREPARING AND QUALIFYING FOR ADMISSION TO TEACHING. 2ND ISSUE OF QUARTERLY SOURCEBOOK, NEW DIRECTIONS FOR EDUCATION. SAN FRANCISCO: JOSSEY-BASS, INC., SUMMER, 1973, 103-125.

COOPER, JAMES M. AND STROUD, THOMAS. "THE STANFORD SUMMER MICRO-TEACHING CLINIC, 1966." MICRO-TEACHING: A DESCRIPTION (MIMEO). STANFORD UNIVERSITY, SCHOOL OF EDUCATION, 1967.

FLANDERS, NED A. TEACHER INFLUENCE, PUPIL ATTITUDES, AND ACHIEVEMENT. UNITED STATES OFFICE OF EDUCATION COOPERATIVE RESEARCH MONOGRAPH NO. 12, OE-25040. WASHINGTON, D.C.: USOE, 1965.

GLASS, GENE V. STATISTICAL AND MEASUREMENT PROBLEMS IN IMPLEMENTING THE STULL ACT. PRESENTATION AT THE STANFORD INVITATIONAL CONFERENCE ON THE STULL ACT, OCTOBER, 1972.

LAWLER, E.S. "DIFFERING RATES OF PROGRESS OF CLASSES UNDER THE SAME AND DIFFERENT TEACHERS." THE JOURNAL OF EDUCATIONAL RESEARCH. 1964, 58, 84-86.

MITZEL, HAROLD E. "TEACHER EFFECTIVENESS." IN CHESTER W. HARRIS (ED.), ENCYLOPEDIA OF EDUCATIONAL RESEARCH. 3RD EDITION. NEW YORK: MCMILLAN, 1960, 1481-1486.

MCNEIL, JOHN D. AND POPHAM, W. JAMES. "THE ASSESSMENT OF TEACHER COMPETENCE." IN R.M.W. TRAVERS (ED.), SECOND HANDBOOK OF RESEARCH IN TEACHING. CHICAGO: RAND MCNALLY, 1973, 218-244.

Popham, W. James. "Applications of Teaching Performance Tests to Inservice and Preservice Teacher Education." Paper presented at the annual meeting of AERA, New Orleans, 1973.

Rosenshine, Barak and Furst, Norma. "Research in Teacher Performance Criteria." In B.O. Smith (Ed.), Research in Teacher Education. Englewood Cliffs, N.J.: Prentice-Hall, 1971, 37-72.

Schalock, H. Del and Ferguson, Jean. "A Guide to Competency Demonstration and Assessment: Extend Full Responsibility Teaching (2-5 weeks). (mimeo) Monmouth, Oregon: Elementary Teacher Education Program, Oregon College of Education. February, 1974. (Field Test Format #1).

Schalock, H. Del. "The Focus: Knowledge, Teaching Behavior, or the Products?" In Joel L. Burdin and Margaret T. Reagan (Eds.), Performance-Based Certification of School Personnel. Washington, D.C.: ERIC Clearinghouse on Teacher Education and the Association of Teacher Educators, 1971, 41-49.

Smith, B.O. "Teaching: Conditions of Its Evaluation." In Pi Lambda Theta, The Evaluation of Teaching. Washington, D.C.: Pi Lambda Theta, 1967, 65-84.

Soar, Robert S. "Accountability: Assessment Problems and Possibilities." Journal of Teacher Education, Fall, 1973, 24, 205, 212.